Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2024 – 07:15 AM EDT

PDB ID	:	5RE1
Title	:	PanDDA analysis group deposition – Endothiapepsin ground state model 58
Authors	:	Weiss, M.S.; Wollenhaupt, J.; Metz, A.; Barthel, T.; Lima, G.M.A.; Heine, A.; Mueller, U.; Klebe, G.
Deposited on	:	2020-03-24
Resolution	:	0.97 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

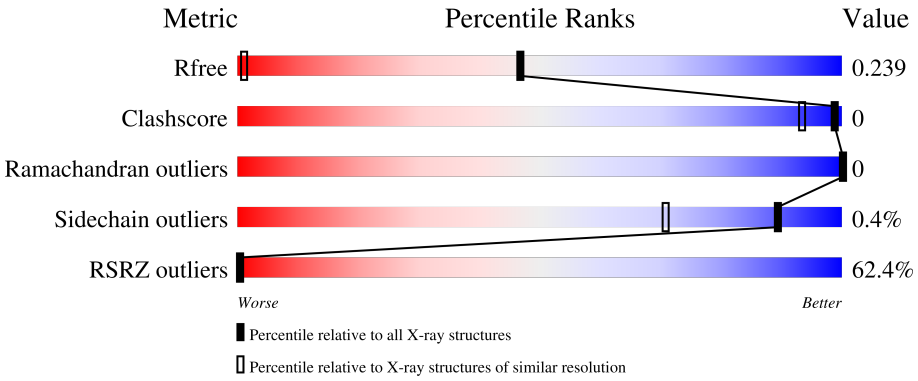
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 0.97 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1026 (1.02-0.94)
Clashscore	180529	1154 (1.02-0.94)
Ramachandran outliers	177936	1094 (1.02-0.94)
Sidechain outliers	177891	1095 (1.02-0.94)
RSRZ outliers	164620	1025 (1.02-0.94)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	419	49% 77% 21%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2578 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Endothiapepsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	330	2462	1566	367	527	2	3	19	0

- Molecule 2 is water.

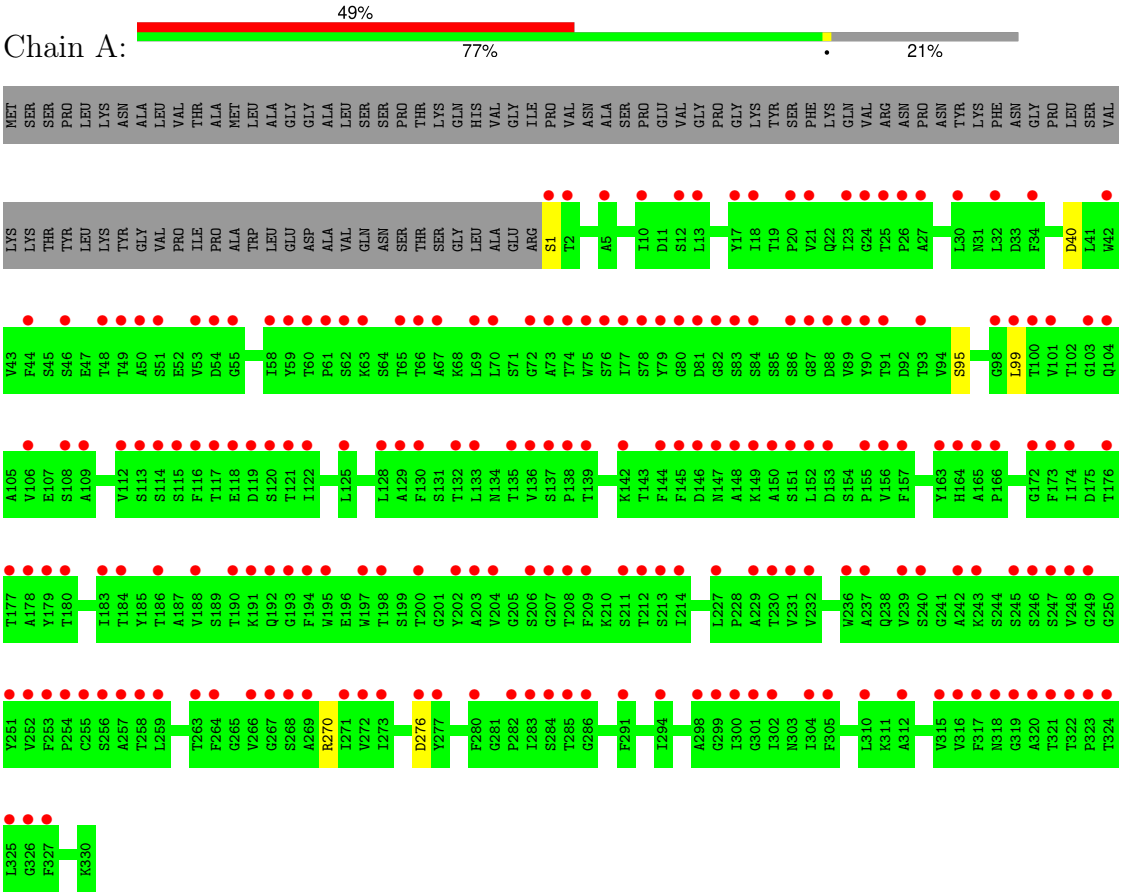
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	116	116	116	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Endothiapepsin

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.18Å 72.85Å 52.42Å 90.00° 109.20° 90.00°	Depositor
Resolution (Å)	49.56 – 0.97 49.51 – 0.97	Depositor EDS
% Data completeness (in resolution range)	91.3 (49.56-0.97) 91.5 (49.51-0.97)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.05 (at 0.97Å)	Xtriage
Refinement program	REFMAC 5.8.0238, PHENIX 1.16.3549	Depositor
R, R_{free}	0.244 , 0.244 0.246 , 0.239	Depositor DCC
R_{free} test set	8737 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	9.2	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 31.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	2578	wwPDB-VP
Average B, all atoms (Å^2)	10.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.62% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.85	1/2552 (0.0%)	0.89	2/3496 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	SER	CA-CB	-5.02	1.45	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	270	ARG	NE-CZ-NH1	8.42	124.51	120.30
1	A	270	ARG	NE-CZ-NH2	-5.37	117.61	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2462	0	2329	2	0
2	A	116	0	0	0	0
All	All	2578	0	2329	2	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 0.

All (2) close contacts within the same asymmetric unit are listed below, sorted by their clash

magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:95:SER:HA	1:A:99:LEU:O	2.20	0.41
1:A:276[B]:ASP:OD1	1:A:276[B]:ASP:N	2.53	0.41

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	347/419 (83%)	344 (99%)	3 (1%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	270/336 (80%)	269 (100%)	1 (0%)	89 71

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	330/419 (78%)	2.67	206 (62%) 0 0	4, 9, 16, 23	19 (5%)

All (206) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	300[A]	ILE	11.3
1	A	321	THR	9.4
1	A	150	ALA	8.6
1	A	248	VAL	8.5
1	A	208	THR	8.1
1	A	149[A]	LYS	7.6
1	A	75	TRP	7.0
1	A	322	THR	6.7
1	A	299[A]	GLY	6.7
1	A	298[A]	ALA	6.6
1	A	50	ALA	6.5
1	A	258	THR	6.2
1	A	320	ALA	6.2
1	A	151	SER	5.7
1	A	247	SER	5.7
1	A	254	PRO	5.7
1	A	25	THR	5.6
1	A	302	ILE	5.6
1	A	115	SER	5.3
1	A	236	TRP	5.2
1	A	209	PHE	5.1
1	A	285	THR	5.1
1	A	259	LEU	5.0
1	A	283	ILE	4.9
1	A	232	VAL	4.9
1	A	291	PHE	4.9
1	A	49	THR	4.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	174[A]	ILE	4.7
1	A	249	GLY	4.6
1	A	26	PRO	4.5
1	A	114	SER	4.5
1	A	246[A]	SER	4.5
1	A	323	PRO	4.5
1	A	190	THR	4.4
1	A	58	ILE	4.4
1	A	122	ILE	4.4
1	A	138	PRO	4.3
1	A	78	SER	4.3
1	A	116	PHE	4.3
1	A	53	VAL	4.3
1	A	74	THR	4.3
1	A	80	GLY	4.3
1	A	51	SER	4.2
1	A	120	SER	4.2
1	A	76	SER	4.2
1	A	13	LEU	4.0
1	A	79	TYR	4.0
1	A	23	ILE	4.0
1	A	177	THR	4.0
1	A	112	VAL	4.0
1	A	113	SER	4.0
1	A	231	VAL	3.9
1	A	90	TYR	3.9
1	A	73	ALA	3.9
1	A	178	ALA	3.9
1	A	89	VAL	3.9
1	A	188	VAL	3.9
1	A	240[A]	SER	3.9
1	A	301	GLY	3.9
1	A	183	ILE	3.9
1	A	42	TRP	3.8
1	A	157	PHE	3.8
1	A	163	TYR	3.8
1	A	317	PHE	3.7
1	A	65	THR	3.7
1	A	93	THR	3.7
1	A	132	THR	3.7
1	A	194	PHE	3.7
1	A	125	LEU	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	133	LEU	3.7
1	A	203	ALA	3.6
1	A	164	HIS	3.6
1	A	176	THR	3.6
1	A	264	PHE	3.6
1	A	84	SER	3.6
1	A	211	SER	3.5
1	A	255	CYS	3.5
1	A	272	VAL	3.5
1	A	55	GLY	3.5
1	A	130	PHE	3.5
1	A	100	THR	3.5
1	A	135	THR	3.5
1	A	63	LYS	3.5
1	A	77	ILE	3.4
1	A	310	LEU	3.4
1	A	229	ALA	3.4
1	A	59	TYR	3.4
1	A	202	TYR	3.3
1	A	198	THR	3.3
1	A	239	VAL	3.3
1	A	252	VAL	3.3
1	A	34	PHE	3.2
1	A	192	GLN	3.2
1	A	165	ALA	3.2
1	A	204	VAL	3.2
1	A	54	ASP	3.2
1	A	81	ASP	3.2
1	A	148	ALA	3.2
1	A	21	VAL	3.2
1	A	101	VAL	3.2
1	A	136	VAL	3.2
1	A	152	LEU	3.2
1	A	271	ILE	3.2
1	A	146	ASP	3.1
1	A	83	SER	3.1
1	A	319	GLY	3.1
1	A	121	THR	3.1
1	A	253	PHE	3.1
1	A	214	ILE	3.1
1	A	137	SER	3.1
1	A	118	GLU	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	256	SER	3.0
1	A	60	THR	3.0
1	A	318	ASN	3.0
1	A	70	LEU	3.0
1	A	82	GLY	3.0
1	A	61	PRO	2.9
1	A	155	PRO	2.9
1	A	327	PHE	2.9
1	A	315	VAL	2.9
1	A	180	THR	2.9
1	A	30	LEU	2.9
1	A	325[A]	LEU	2.9
1	A	67	ALA	2.9
1	A	173	PHE	2.9
1	A	273	ILE	2.8
1	A	86	SER	2.8
1	A	87	GLY	2.8
1	A	103	GLY	2.8
1	A	212	THR	2.8
1	A	266	VAL	2.8
1	A	197	TRP	2.8
1	A	72	GLY	2.8
1	A	257	ALA	2.8
1	A	69	LEU	2.8
1	A	191	LYS	2.8
1	A	106	VAL	2.8
1	A	2	THR	2.7
1	A	12	SER	2.7
1	A	17	TYR	2.7
1	A	179	TYR	2.7
1	A	186	THR	2.7
1	A	213	SER	2.7
1	A	144	PHE	2.6
1	A	27	ALA	2.6
1	A	66	THR	2.6
1	A	263	THR	2.6
1	A	153	ASP	2.6
1	A	206[A]	SER	2.6
1	A	267	GLY	2.6
1	A	195	TRP	2.5
1	A	294	ILE	2.5
1	A	24	GLY	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	286	GLY	2.5
1	A	128	LEU	2.5
1	A	147	ASN	2.5
1	A	145	PHE	2.5
1	A	98	GLY	2.5
1	A	304	ILE	2.5
1	A	251	TYR	2.4
1	A	5	ALA	2.4
1	A	109	ALA	2.4
1	A	269	ALA	2.4
1	A	193	GLY	2.4
1	A	18	ILE	2.4
1	A	156	VAL	2.4
1	A	117	THR	2.4
1	A	139	THR	2.4
1	A	46	SER	2.4
1	A	44	PHE	2.4
1	A	243	LYS	2.4
1	A	245	SER	2.4
1	A	20	PRO	2.3
1	A	280	PHE	2.3
1	A	326	GLY	2.3
1	A	48	THR	2.3
1	A	324	THR	2.3
1	A	62	SER	2.3
1	A	10	ILE	2.3
1	A	316	VAL	2.3
1	A	119	ASP	2.3
1	A	237	ALA	2.3
1	A	32	LEU	2.3
1	A	108	SER	2.3
1	A	268[A]	SER	2.3
1	A	200	THR	2.2
1	A	172	GLY	2.2
1	A	282	PRO	2.2
1	A	1	SER	2.2
1	A	312	ALA	2.2
1	A	99	LEU	2.2
1	A	142	LYS	2.2
1	A	184	THR	2.2
1	A	104	GLN	2.2
1	A	242	ALA	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	227	LEU	2.1
1	A	88	ASP	2.1
1	A	284	SER	2.1
1	A	91	THR	2.1
1	A	207	GLY	2.1
1	A	230	THR	2.1
1	A	277	TYR	2.1
1	A	166	PRO	2.1
1	A	276[A]	ASP	2.1
1	A	129	ALA	2.0
1	A	305	PHE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.